

Online Convex Optimization (FTRL, OMD)

Lecturer: Kris Kitani

Scribes: Haowen Shi, Fan Jia

1 Review

In the last lecture, we studied a generalized framework of online learning algorithms called *Online Convex Optimization* (OCO) [2], as shown in Fig. 1.

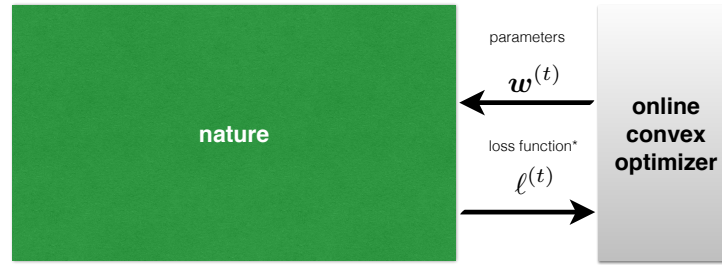


Figure 1: OCO framework

We introduced our first algorithm called *Follow the Leader* (FTL) known as ‘Fictitious Play’ (Brown 1951 [1]). The prediction rule is that the learner is going to choose the parameter \mathbf{w} such that it minimizes the cumulative loss up to now. This rule can be inserted into the online convex optimizer in Fig. 1.

Algorithm 1 FTL

```

1: function FOLLOW THE LEADER
2:   for  $t = 1, 2, \dots, T$  do
3:      $\mathbf{w}^{(t)} = \arg \min_{\mathbf{w} \in W} \sum_{i=1}^{t-1} f^{(i)}(\mathbf{w})$  ▷ Parameter chosen
4:     RECEIVE  $(f^{(t)} : W \rightarrow \mathbb{R})$  ▷ Receive loss function
5:   end for
6: end function
  
```

We used “the one-step look ahead cheater” to upper bound the regret:

$$R(\mathbf{u}) = \sum_t [f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{u})] \leq \sum_t [f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)})]$$

In this lecture, we study the FTL algorithm with different loss functions. We will see how linear loss function will make the FTL unstable and then introduce regularization to resolve such instability.

2 Types of Loss Functions

In the “Follow the leader” algorithm, the environment returns a loss function. Loss functions affect the algorithm’s performance guarantees. Fig. 2 shows some examples of loss functions.

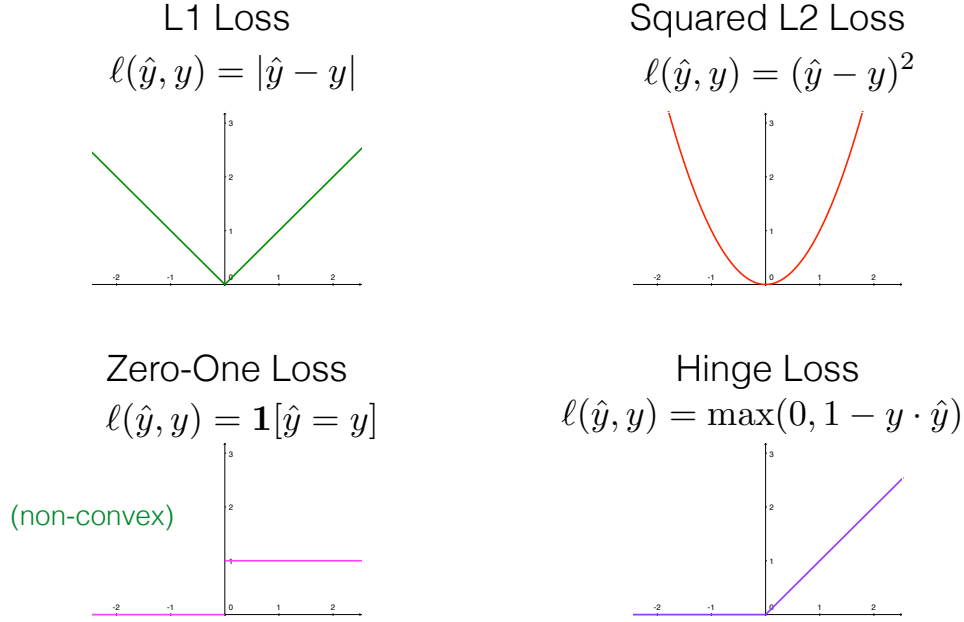


Figure 2: Many types of loss function

In section 3, we take a look at FTL with a quadratic (squared L2) loss function as a concrete example.

3 Follow The Leader with Quadratic Loss

Algorithm 2 FTL with Quadratic Loss

```

1: function FOLLOW THE LEADER (QUADRATIC LOSS)
2:    $f^{(0)} \leftarrow 0$ 
3:   for  $t = 1, 2, \dots, T$  do
4:      $\mathbf{w}^{(t)} = \arg \min_{\mathbf{w} \in S} \sum_{i=1}^{t-1} f^{(i)}(\mathbf{w})$ 
5:     RECEIVE ( $f^{(t)} = \frac{1}{2} \|\mathbf{w} - \mathbf{z}^{(t)}\|_2^2$ ) ▷ Receive loss function
6:   end for
7: end function

```

In this algorithm, the entire loss function comes from the environment. \mathbf{z} is a parameter of the loss function and \mathbf{w} is the input. Since the average of \mathbf{z} we have seen up to now is the minimizer of the

quadratic loss, we can rewrite step 4 as:

$$\mathbf{w}^{(t)} = \frac{1}{t-1} \sum_{i=1}^{t-1} z^{(i)}$$

3.1 Regret Bound

Assume that nature always chooses a quadratic loss:

$$f^{(t)}(w) = \frac{1}{2} \|w - z^{(t)}\|_2^2$$

We assume that points are bounded:

$$\|z\|_2^2 \leq L$$

Recall that our generic FTL regret bound is:

$$R(\mathbf{u}) = \sum_t [f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{u})] \leq \sum_t [f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)})]$$

Now we inset the quadratic loss into the right hand side of the bound:

$$\begin{aligned} f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)}) &= \frac{1}{2} \|w^{(t)} - z^{(t)}\|_2^2 - \frac{1}{2} \|w^{(t+1)} - z^{(t)}\|_2^2 \\ &= \frac{1}{2} \|w^{(t)} - z^{(t)}\|_2^2 - \frac{1}{2} \|(1 - \frac{1}{t})w^{(t)} + (\frac{1}{t})z^{(t)} - z^{(t)}\|_2^2 \end{aligned}$$

Recall that mean of z is the quadratic loss minimizer:

$$w^{(t)} = \frac{1}{t-1} \sum_{i=1}^{t-1} z^{(i)}$$

We can convert it in a incremental, then we have the update step:

$$w^{(t+1)} = (1 - \frac{1}{t})w^{(t)} + (\frac{1}{t})z^{(t)}$$

Now we plug in minimizer and upper-bound it:

$$\begin{aligned} f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)}) &= \frac{1}{2} \|w^{(t)} - z^{(t)}\|_2^2 - \frac{1}{2} \|(1 - \frac{1}{t})w^{(t)} + (\frac{1}{t})z^{(t)} - z^{(t)}\|_2^2 \\ &= \frac{1}{2} \left(1 - (1 - \frac{1}{t})^2\right) \|w^{(t)} - z^{(t)}\|_2^2 \end{aligned}$$

Use the ‘known’ inequality:

$$\frac{1}{2} \left(1 - (1 - \frac{1}{t})^2\right) \leq \frac{1}{t}$$

We can have:

$$f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)}) \leq \left(\frac{1}{t}\right) \|w^{(t)} - z^{(t)}\|_2^2$$

Recall our prior assumptions:

$$L = \max_t \|z^{(t)}\|$$

Since average is less or equal to max, we have:

$$\|w^{(t)}\| \leq L$$

Based on triangle inequality, we can get:

$$\|w^{(t)} - z^{(t)}\| \leq 2L$$

Therefore, we have:

$$f^{(t)}(w^{(t)}) - f^{(t)}(w^{(t+1)}) \leq \left(\frac{1}{t}\right) 4L^2$$

Recall the regret bound over all time steps is:

$$R(\mathbf{u}) \leq \sum_t [f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)})]$$

We know another useful approximation:

$$\sum_{t=1}^T \leq 1 + \int_1^T (1/t) dt = \log(T) + 1$$

Therefore, the regret bound is:

$$\text{Regret} \leq 4L^2 (\log(T) + 1)$$

This is no-regret since the average regret grows sub-linearly in T. However, this only holds for quadratic loss functions, not for all convex functions.

4 Follow The Leader with Linear Loss

Algorithm 3 Follow the leader algorithm with linear loss

```

1: function FOLLOW THE LEADER (LINEAR LOSS)
2:   for  $t = 1, 2, \dots, T$  do
3:      $\mathbf{w}^{(t)} = \arg \min_{\mathbf{w} \in S} (\sum_{i=1}^{t-1} \mathbf{z}^{(i)}) \cdot \mathbf{w}$ 
4:     RECEIVE  $(f^{(t)}(\mathbf{w}) = \mathbf{z}^{(t)} \cdot \mathbf{w})$  ▷ Receive loss function
5:   end for
6: end function

```

In this algorithm, the space of possible predictions is $\mathbf{w} \in S = [-1, 1]$. The nature always chooses

a linear loss: $f^{(t)}(\mathbf{w}) = \mathbf{z}^{(t)} \cdot \mathbf{w}$, where $\mathbf{z}^{(t)} = \begin{cases} -0.5 & \text{if } t \text{ is 1} \\ 1 & \text{if } t \text{ is even} \\ -1 & \text{if } t \text{ is odd} \end{cases}$

The total loss of this FTL learner is:

$$L = \sum_{t=1}^T f^{(t)} = \sum_{t=2}^T 1 = T - 1$$

The total loss of expert $u = 0$:

$$L_u = \sum_{t=2}^T z^{(t)} \cdot u = \sum_{t=2}^T 0 = 0$$

The regret of FTL with linear loss relative to expert u that chooses $w = 0$:

$$\text{Regret}(u) = T - 1 - 0 = T - 1 = O(T)$$

The regret grows linearly. It is because the \mathbf{w} value flip-flops back and forth to make the algorithm unstable. To resolve this issue, we will introduce regularization in next section.

5 Follow The Regularized Leader

In section 4 we saw that FTL with linear loss is very unstable. To make \mathbf{w} more stable, here we introduce the *Follow the Regularized Leader* (FTRL) algorithm.

Algorithm 4 Follow The Regularized Leader

```

1: function FTRL(CONVEX SET  $\mathcal{S}$ )
2:   for  $t = 1, 2, \dots, T$  do
3:      $\mathbf{w}^{(t)} = \arg \min_{\mathbf{w} \in \mathcal{S}} \sum_{i=1}^{t-1} f^{(i)}(\mathbf{w}) + \psi(\mathbf{w})$  ▷ Added  $\psi(\mathbf{w})$  to the prediction rule
4:     RECEIVE ( $f^{(t)} : \mathcal{S} \rightarrow \mathbb{R}$ )
5:   end for
6: end function

```

Note that FTL is just FTRL with $\psi = 0$.

5.1 Regret Bound

The regret bound for FTRL is similar to that of the FTL, with an additional term due to regularization:

$$\begin{aligned}
R(\mathbf{u}) &\triangleq \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{u}) \right] \\
&\leq \left[\psi(\mathbf{u}) - \psi(\mathbf{w}^{(1)}) \right] + \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)}) \right]
\end{aligned}$$

Proof. Given the regret bound for FTL using one-step look ahead cheater:

$$\begin{aligned} R(\mathbf{u}) &= \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{u}) \right] \\ &\leq \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)}) \right] \end{aligned} \quad (1)$$

Add the $t = 0$ 'loss' for FTRL on both sides of (1):

$$\begin{aligned} &\left[f^{(0)}(\mathbf{w}^{(0)}) - f^{(0)}(\mathbf{u}) \right] + \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{u}) \right] \\ &\leq \left[f^{(0)}(\mathbf{w}^{(0)}) - f^{(0)}(\mathbf{u}) \right] + \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)}) \right] \end{aligned} \quad (2)$$

Swap in the regularizer notation:

$$\begin{aligned} &\left[\phi(\mathbf{w}^{(0)}) - \phi(\mathbf{u}) \right] + \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{u}) \right] \\ &\leq \left[\phi(\mathbf{w}^{(0)}) - \phi(\mathbf{u}) \right] + \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)}) \right] \end{aligned} \quad (3)$$

Subtract $\phi(\mathbf{w}^{(0)}) - \phi(\mathbf{u})$ from both sides:

$$\sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{u}) \right] \leq \left[\psi(\mathbf{u}) - \psi(\mathbf{w}^{(1)}) \right] + \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)}) \right] \quad (4)$$

□

5.2 FTRL with Linear Loss

Now we revisit the linear loss function scenario and see if FTRL can make \mathbf{w} more stable using quadratic regulation. We use a squared L2 norm as a quadratic regularizer:

$$\phi(\mathbf{w}) = \frac{1}{2\eta} \|\mathbf{w}\|_2^2 \quad (5)$$

The linear loss function is the same as before:

$$f^{(t)} = \mathbf{w} \cdot \mathbf{z}^{(t)} \quad (6)$$

With quadratic regularization, the prediction rule now becomes an unconstrained quadratic programming problem. For convenience of analysis, we introduce the following short-hand for the sum:

$$\theta^{(t)} = \sum_{i=1}^{t-1} \mathbf{z}^{(i)} \quad (7)$$

Algorithm 5 FTRL with Linear Loss And Quadratic Regulation

```
1: function FTRL(CONVEX SET  $\mathcal{S}$ )
2:   for  $t = 1, 2, \dots, T$  do
3:      $\mathbf{w}^{(t)} = \arg \min_{\mathbf{w} \in \mathcal{S}} \sum_{i=1}^{t-1} \left( \frac{1}{2\eta} \|\mathbf{w}\|_2^2 + \mathbf{w} \cdot \sum_{i=1}^{t-1} z^{(i)} \right)$  ▷ Prediction rule
4:     RECEIVE  $(f^{(t)} : \mathbf{w} \cdot z^{(t)})$ 
5:   end for
6: end function
```

We will later see that $\theta^{(t)}$ is the sum of loss function gradients. With this short-hand, we can write the prediction rule as:

$$\mathbf{w}^{(t)} = \arg \min_{\mathbf{w} \in \mathcal{S}} \sum_{i=1}^{t-1} \left(\frac{1}{2\eta} \|\mathbf{w}\|_2^2 + \mathbf{w} \cdot \theta^{(t)} \right) \quad (8)$$

To minimize the quadratic function in (8), we calculate its first derivative and set to zero:

$$\begin{aligned} \frac{d}{d\mathbf{w}} \left(\frac{1}{2\eta} \|\mathbf{w}\|_2^2 + \mathbf{w} \cdot \theta^{(t)} \right) &= 0 \\ \frac{1}{2\eta} 2\mathbf{w} + \theta^{(t)} &= 0 \\ \mathbf{w} &= -\eta \theta^{(t)} \end{aligned}$$

Substitute (7):

$$\begin{aligned} \mathbf{w}^{(t)} &= -\eta \sum_{i=1}^{t-1} z^{(i)} \\ &= -\eta \left(\sum_{i=1}^{t-2} z^{(i)} + z^{(t-1)} \right) \\ &= -\eta \sum_{i=1}^{t-2} z^{(i)} - \eta z^{(t-1)} \\ &= \mathbf{w}^{(t-1)} - \eta z^{(t-1)} \end{aligned} \quad (9)$$

Note that (9) is an update equation for \mathbf{w} across each iteration. The hyper parameter η controls the step size / learning rate. This convenience of doing optimization with incremental updates is really due to the fact that we chose a good regularizer function. Otherwise we would have had to solve a full optimization problem in each iteration.

Also note that we have seen a similar form of (9) before in the perceptron algorithm. We will see this update form again in future gradient descent algorithms and the general form of this update equation is called online gradient descent:

$$\mathbf{w}^{(t)} = \mathbf{w}^{(t-1)} - \eta \nabla f^{(t-1)} \quad (10)$$

5.3 Regret of FTRL with Euclidean Regularizer and Linear Loss

The regret bound for FTRL with a Euclidean regularizer and linear loss function is:

$$R^{(T)}(\mathbf{u}) \leq BL\sqrt{2T} \quad (11)$$

Where $L = \max_{\mathbf{z}} \|\mathbf{z}\|_2$ and $B = \max_{\mathbf{u} \in \mathcal{S}} \|\mathbf{u}\|_2$

Proof.

$$\begin{aligned}
R^{(T)}(\mathbf{u}) &\leq \left[\phi(\mathbf{u}) - \phi(\mathbf{w}^{(1)}) \right] + \sum_{t=1}^T \left[f^{(t)}(\mathbf{w}^{(t)}) - f^{(t)}(\mathbf{w}^{(t+1)}) \right] && \text{Generic form} \\
&= \left[\frac{1}{2\eta} \|\mathbf{u}\|_2^2 - 0 \right] + \sum_{t=1}^T \left[\langle \mathbf{w}^{(t)}, \mathbf{z}^{(t)} \rangle - \langle \mathbf{w}^{(t+1)}, \mathbf{z}^{(t)} \rangle \right] && \text{Plug in regularizer} \\
&= \frac{1}{2\eta} \|\mathbf{u}\|_2^2 + \sum_{t=1}^T \langle \mathbf{w}^{(t)} - \mathbf{w}^{(t+1)}, \mathbf{z}^{(t)} \rangle && \text{Rearrangement} \\
&= \frac{1}{2\eta} \|\mathbf{u}\|_2^2 + \sum_{t=1}^T \eta \|\mathbf{z}^{(t)}\|^2
\end{aligned}$$

So we have the following regret bound:

$$\begin{aligned}
R^{(T)}(\mathbf{u}) &\leq \frac{1}{2\eta} \|\mathbf{u}\|_2^2 + \sum_{t=1}^T \eta \|\mathbf{z}^{(t)}\|^2 \\
R^{(T)}(\mathbf{u}) &\leq \frac{1}{2\eta} B^2 + \eta T L^2 \\
R^{(T)}(\mathbf{u}) &\leq B L \sqrt{2T} && \text{Optimal value at } \eta = \frac{B}{L\sqrt{2T}}
\end{aligned}$$

□

Because the regret bound (11) is sub-linear, FTRL with an Euclidean regularizer and linear loss is a no regret algorithm.

Now that we have studied the FTRL framework, looking back at the online perceptron algorithm and RWMA, their update rules actually come from the general form of FTRL update rule (10). This is illustrated in Fig. 3 Specifically, both online perceptron algorithm and RWMA have linear loss. Their difference is that the online perceptron algorithm uses an Euclidean regularization but the RWMA uses an Entropic regularization.

6 Online Mirror Descent

Online Mirror Descent (OMD) can be considered as a special case of FTRL with a linear loss and a convex regularizer. Because we did not finish discussing OMD in the lecture, details on OMD will be covered in our next scribe notes.

Online Gradient Descent
(Online Perceptron Algorithms)

$$f^{(t)}(\mathbf{w}) = \mathbf{w} \cdot \mathbf{x}^{(t)}$$

linear predictor (loss)

$$(\mathbf{w}) = \frac{1}{\eta} \|\mathbf{w}\|_2^2$$

Euclidean regularization

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} - \eta \nabla \ell(y^{(t)}, \hat{y}^{(t)})$$

Additive update

$$BL\sqrt{2T}$$

No Regret

Online Exponentiated Gradient Descent
(Hedge/GWM, Winnow)

$$f^{(t)}(\mathbf{w}) = \mathbf{w} \cdot \mathbf{x}^{(t)}$$

linear predictor (loss)

$$(\mathbf{w}) = \frac{1}{\eta} \sum_n w_n \log(w_n)$$

Entropic Regularization

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} \exp\left(-\eta \nabla \ell(y^{(t)}, \hat{y}^{(t)})\right)$$

Exponential update

$$\sqrt{(T/2) \log N}$$

No Regret

Figure 3: FTRL examples that we have seen before

References

- [1] U. Berger. Brown's original fictitious play. *Journal of Economic Theory*, 135(1):572–578, 2007.
- [2] S. Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and trends in Machine Learning*, 4(2):107–194, 2011.